

USING EMOTIONAL STROOP TASK AS AN AUXILIARY TOOL IN DIAGNOSIS OF PAEDOPHILIA¹

Krzysztof KASPAREK

Centre for Evaluation and Analysis of Public Policies, Jagiellonian University, Kraków, Poland

Summary

This work concerns an issue of using results of emotional Stroop task as an indicator facilitating diagnosis of sexual preferences of paedophilia type. Publications whose authors, basing on indicators created on the basis of emotional Stroop task tried to select persons manifesting sexual interest in children, are discussed here. Further one presents results of own research using Stroop task, conducted on 26 years old man manifesting non-exclusive paedophilic disorder. On the basis of the used procedure one obtained results which can be interpreted as concordant with a diagnosis of non-exclusive paedophilia.

Key words

Paedophilia; Diagnosis; Assessment; Stroop.

Received 1 July 2016; accepted 27 July 2016

1. Introduction

Stroop task (Stroop, 1935) was designed as one of task used for measurement of cognitive processes. Examined person's task is to properly categorise colours in which are written words of various semantic meaning. Reaction time between presentation of a given word and assigning it a colour by an examined person is recognised as an indicator of interference that occurs between a word semantic meaning and an automatized process of naming a colour. It is assumed that differences in reaction times needed to categorise a colour of presented word categories are caused by a difference in their importance/meaning for an examined person. Reactions may be based both on pressing a button representing a given colour, as well as on a loud uttering of a name of a given colour.

Among the most popular modifications of Stroop task P. Smith (2009) specifies:

- words and colours – examined person is presented with various categories of words, to which s/he is supposed to assign colours independently of their semantic meaning;
- words and pictures – examined person is presented with pictures with words placed on them. Examined person's task is to categorise pictures regardless of words' semantic meaning;
- sounds and words – examined person's task is to classify presented sounds as high and low, regardless of the semantic meaning of a sound (e.g. word "low" presented with a high sound tone);
- global-local – examined person's task is to categorise a global (general) shape of stimuli, regardless of how they are presented (e.g. number five composed from small numbers seven).

Another, very often occurring type of modification is so called emotional Stroop. It consists in supplementation of task with additional variable – emotional valence of presented stimuli. Stimuli that require

¹ The present paper is a fragment of unpublished master's thesis on *Possibilities to use indicators based on reaction time in diagnosis of paedophilic disorder based on an example of emotional Stroop task*, written under a supervision of Professor J. K. Gierowski, presented at the Institute of Psychology, Jagiellonian University.

proper categorisation of colour represent a semantic meaning related to emotion of sadness, happiness, anger, etc. The described type of modification may also contain stimuli related to dysfunctions manifested by examined persons from experimental group. For example, in a study in which participate persons who are addicted to alcohol, one of stimuli groups directly refers to alcohol (e.g. a word “glass”, picture of a bottle of beer, and so on).

A version of Stroop task, modified in the described way, was used in studies in which participated persons suffering from, among others, anxiety disorders (MacLeod, Matthews, Tata, 1986, in: Smith, 2009), depression (Gotlib, McCann, 1984, in: Smith, 2009), PTSD (Foa, Feske, Murdock, Kozak, McCarthy, 1991, in: Smith, 2009), social phobia (Hope, Rapee, Heimberg, Dombeck, 1990, in: Smith, 2009), as well as substance addicts (Banich et al., 2007, in: Price, Beech, Mitchel, Humphreys, 2013). In the light of the presented research results persons suffering from the above-mentioned disorders needed significantly more time for categorising colours of words which were related to disorders they were manifesting, when compared to control groups. In other words, stimuli originating from categories that were problematic for examined populations required longer reaction times (*RT*) than stimuli from neutral categories, when compared to persons from control groups (Price, Hanson, 2007).

2. Stroop task in examination of child sexual offenders

Results of reaction elongation to stimuli related to manifested dysfunctions, observed in cognitive tasks, became an inspiration to researchers of sexual offenders. One detected here both a chance to investigate thoroughly a mechanism that pushes a person to commit the mentioned crimes, as well as a possibility to constitute an objective, less invasive, not easy to manipulate, cheaper in exploitation, indicator of disturbed sexual preferences. As one of works which authors initiated this approach, can be found a paper by P. Smith and M. Waterman (2004), which was published in the *Sexual Abuse: A Journal of Research and Treatment*. The authors examined child sexual offenders with the procedure of modified Stroop task in emotional version. As a control group they selected sexual offenders who had attacked women, offenders of aggressive, non-sexual crimes, offenders of non-aggressive crimes, and students of the Leeds University.

Price et al. (2013), who described the most important of obtained results, indicate:

- relatively longer average *RT* in case of words with sexual meaning among offenders of sexual crimes, when compared to other groups;
- relatively longer average *RT* in case of words from aggressive category among aggressive offenders. Perpetrators who committed rapes in this condition needed relatively more time than offenders of less aggressive crimes;
- lack of significant differences in presented *RT* between child sexual offenders and sexual offenders who attacked adults.

Summarising their study, Smith and Waterman (2004) concluded that words used in Stroop task that were close in their meaning to the character of crimes committed by examined offenders, activate their cognitive associations that are absent in other examinees. According to authors the typical pattern of relatively longer *RTs* reflects exactly this process.

The bold hypothesis and promising results caused a real flood of research on use of various cognitive tasks in examination of sexual offenders. The Stroop emotional task also waited until it was continued. Its replications by particular authors with participation of child sexual offenders are compiled in Table 1².

Study by Price and Hanson (2007), which was a partial replication of the described experiment by Smith and Waterman, gave similar results. An average *RT*, required by rape offenders to categorise colours of words with sexual meaning, were significantly longer than in neutral condition, when compared to persons who were not sentenced for any crimes. A special attention deserves a fact that the authors in their experiment used both Stroop task with the battery of words proposed by Smith and Waterman, as well as their own set of words with emotional meaning. In case of second set one did not observe statistically significant differences in time of response. Authors explain that with an “intuitive” method of words selection they adopted. A list of categories of words used in the discussed research and in other analysed experiments is presented in Table 2.

² On purpose one resigned from including to this list a study by Pflugradt and Allen (2010). It concerned female sexual offenders, many of whom committed sexual abuse of a child, but Stroop task used in this study was used only in classical version, and it aimed at diagnosis of cognitive control efficiency. In case of the second study, a sample of sexual offenders constituted only juvenile offenders, and information on them made it impossible to determine if there were child sexual offenders among them.

Table 1
Overview of studies on Emotional Stroop Task as an indirect measure of deviant sexual interest in children

Research team	Stroop task version	Stimulus category	Participants (n)	Control (n)
Smith & Waterman (2004)	ES	Words sexual vs. neutral	CSA(5)	RPE(5) VO(10) NVO(10) NO(13) ^{PS}
Price & Hanson (2007)	ES	Words sexual vs. neutral ^{ZA}	CSA(15)	RPE(15), VO(15), NVO(15), NO(15)
Ó Ciardha & Gormley (2012)	PS	Pictures adults vs. children	CSA (24) 14 incarcerated	NO(24) ^{ZP}
Van Leeuwen et al. (2013)	P-WS	Pictures + words adults vs. children	PED(19) community sample	NO (19) ^{n/a}
Price et al. (2013)	ES	Words sexual vs. neutral ^{ZA}	ICT(4), PED(9), MIX(2), EXH(5), RPE(7),	VO(21), NO(38) ^{PSP}

ES – emotional Stroop task, PS – pictorial Stroop task, P-WS – picture-word Stroop; ^{ZA} – Smith & Waterman (2004) stimulus set + stimulus set design by authors; CSA – child sexual abuser, ICT – incest abuser, PED – paedophiles, MIX – offenders with mixed types of victims (adults + children), EXH – exhibitionists, RPE – rapists; VO – violent, nonsexual offenders, NVO – nonviolent offenders, NO – Non-offender controls; ^{SP} – psychology students, ^{ZP} – males interested in psychology, ^{n/a} – no data.

Data adapted from Schmidt et al. (2015).

Table 2
Stimulus types used in studies on emotional Stroop task as an indirect measure of deviant sexual interest in children

Smith & Waterman (2004)	Price & Hanson (2007)	Price et al. (2013)	Ó Ciardha & Gormley (2012)	van Leeuwen et al. (2013)
Neutral ^W	Neutral ^W	Emotional-personality descriptors ^W	Adult males ^P	Neutral ^W
Aggressive ^W	Child molesting ^W	Sexual actions ^W	Adult females ^P	Sexual ^W
Positive emotions ^W	Sex/intimacy ^W	Physical descriptors ^W	Children (boys) ^P	Adult males ^P
Negative emotions ^W	Rape ^W		Children (girls) ^P	Adult females ^P
Sexual ^W			Neutral (cats) ^P	Children (boys) ^P
Colour ^W			Colour ^W	Children (grils) ^P

^W – words; ^P – pictures.

Another important study where emotional Stroop task was used was experiment by Ó Ciardha and Gormley (2012). Examined persons' task was to properly categorise colours of pictures presenting figures of adult women and men, children and cats in control task. It is the only one among discussed studies where authors took into account homosexual persons both among child sexual offenders, and among persons from control group. Due to applied intervention it was possible to select pictures concordant with sexual interests of examined persons (e.g. adult women). One of the most interesting results obtained in the discussed experiment is statistically important relative difference in *RT* in case of pictures of children whose sex was concordant with sexual interests of an offender, when compared to pictures of children whose sex was dis-

cordant with offender's interests. It is worthwhile to emphasise the fact that this effect was observed only in a case of offenders who openly admitted their preference for children in sexual contacts.

In research by van Leeuwen et al. (2013) Stroop task was used in word-picture version. Examined persons were classifying words placed on pictures that were displayed to them as sexual or neutral. Pictures, similarly like in the study described above, presented figures of both sexes, both children and adults. In the presented variant of Stroop task, named by the author as Picture Association Task³ one obtained an effect that was opposite than in other studies. Persons who

³ The author of present paper, following the review by Schmidt, Banse and Imhoff (2015), decided to classify this task as a modified Stroop task.

self-identified as having sexual interests in the form of paedophilia obtained significantly shorter *RTs* in situations, when a sexual word was placed on a picture of a small child than in other conditions, when compared to control group. It is worthwhile to pay attention to the fact that in the cited research one conducted separate analyses for persons being hetero- and homosexual paedophiles.

The last of the cited research (Price, Beech, Mitchell, Humphreys, 2013), similarly like experiment by Price and Hanson (2007), was composed of replication of procedure by Smith and Waterman (2004) and a new set of test words. One of its most important characteristics was a method used to select words with sexual meaning. On the basis of sexual offenders' answers to one of open questions from the Relapse Prevention Questionnaire (Beckett, Fisher, Mann, Thornton; 1997, in: Price, Beech, Mitchell, Humphreys, 2013), one selected terms describing sexual activities which appeared most often in their utterances. Similarly like in study by Price and Hanson (2007), Stroop task based on Smith's and Waterman's set of words proved averagely longer *RT* needed for answering in sexual condition than in neutral condition in case of sexual offenders, when compared to persons who were not sentenced. The same effect was observed among sexual offenders whose victims were exclusively adult persons, when compared to aggressive offenders. One did not manage to prove the mentioned differences between the totality of sexual offenders and aggressive offenders.

3. Examined person

The examined person was G. – 26 years old man, graduate of IT college. After graduating his school, G. undertook unsuccessful attempts to start university studies or to work as a blue-collar. He gave up each attempt because of difficulties he had in social adaptation and anxiety he experienced because of that. During his life he had not had close relationships with peers. Also, he had never had an intimate relationship. G. participated in an individual therapy which he started voluntarily because of his sexual interest in children. He had noticed his sexual preferences in the form of paedophilia when he was adolescent and was about 17–18 years old. More or less since that time he had fantasised about having sexual relationships with boys aged 10–12. Apart from fantasising, G. consumed also pornographic materials presenting boys in the mentioned age. When talking with his therapist, G. denied him ever having sexual contacts with minors. He also

had not had any sexual contacts with sexually mature persons. Apart from child pornography, G. consumed also pornography with adults. As sexually attractive he described subtle, shy women.

After a conversation with his therapist and a short description of the aim of study G. agreed to participate. During his contact with researcher, after recalling the aim of research and short presentation of procedure, G. maintained his will to participate in research.

4. Research procedure

Applied procedure was based on experiment by Smith and Waterman (2004), in which category of sexually marked words was replaced by “sexual abuse of children” category from experiment by Price and Hanson (2007). The final set was composed of 150 words grouped in seven thematic categories: neutral, positive, related to aggression, positive emotions, negative emotions, colours, and words related to sexual abuse of children. Words used in the procedure together with their original version are presented in the Appendix.

Research procedure was programmed in DMDX software on Epson E-Machines E525z laptop, with 15.60” screen, 2GHz processor, and Windows XP operating system. Reactions of the examined person consisted in pressing one of keys (A, S, K, L) corresponding with a colour of presented word (red, green, yellow, or blue). To make proper answering easier for the examinee, colourful markers were stuck on the keys, which were in a colour corresponding with the one represented by the key. It is worthwhile to observe that this procedure differs from the original experiments, where examined persons' task was to utter loudly the name of presented word's colour.

The procedure started with instruction, and after learning this the examined person had a training session composed of 29 tests. During the training he received feedback concerning correctness or lack of correctness of his reactions. Each of tests started with 500 ms long presentation of “X” fixation point in the centre of the screen, and afterwards one of 150 words was randomly displayed. Each of them could be displayed only once during a procedure. Information on the end of task was presented after an examined person reacted to the last of presented words.

5. Analysis of results

Similarly like in original study by Smith and Waterman (2004) and Price and Hanson (2007), in the

conducted research differences in *RT* in emotional conditions compared to *RT* in neutral conditions were analysed. Additional modification, absent in the described experiments, consisted in separation of ten words concerning general sexual behaviours and intimacy, not related directly to children, from category of words related to child sexual abuse. This separation was conducted according to suggestion from that experiment (Price, Hanson, 2007).

6. Exclusion from analyses

Despite the previous testing of Stroop task, in the original version of this task were small mistakes consisting in improper displaying of five words, among which two were from aggression category, two from negative emotions category and one from colours category. In case of four of them errors consisted in a lack of Polish diacritics (ą, ć, ł). The last from problematic words was displayed with a space between letters (“aggre_ssive”). To level the influence of wrongful way of displaying the words on examined person’s *RT* they were excluded from further analyses.

Similarly like in original research, times of erroneous reactions (improper categorisation of colour) were not included to analyses. From analyses one also

excluded outliers. Problems that were related to that together with a scheme of research will be discussed in more details later.

7. Results of research

During procedure the examined person made one mistake in categorisation of colours in a word from neutral category. The result was deleted from further analyses. A participant’s *RT* spread chart divided to word categories was presented on Figure 1.

On the basis of analysis of the presented chart one decided to exclude *RTs* being beyond 5 and 95 percentile, namely:

- 3 observations from neutral category;
- 1 observation from positive emotions category;
- 1 observation from negative emotions category;
- 1 observation from colours category.

Excluded observations constituted 8% of all registered *RTs*. Descriptive statistics of such prepared *RTs* are presented in Table 3. Category of words taken from research by Price and Hanson (2007) was presented both collectively, as well as after breaking into subcategories described above, that were related to sexual actions towards children and generally under-

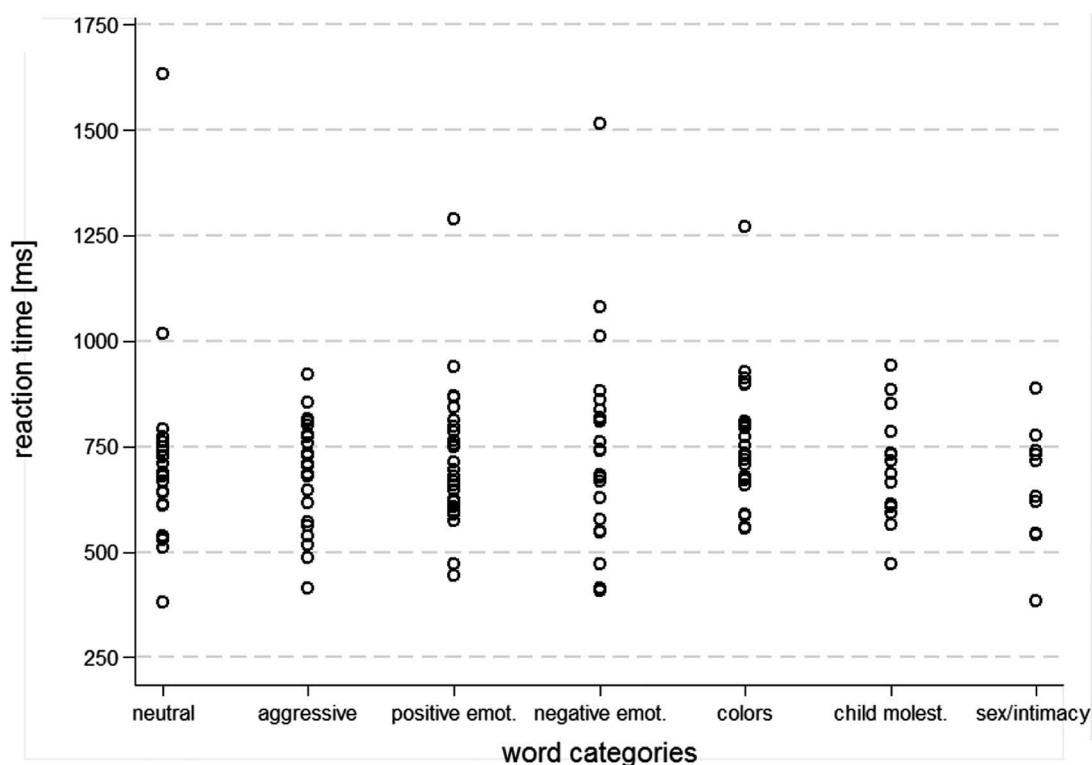


Fig. 1. Participant’s reaction time.

stood sexuality and intimacy (without participation of children).

The longest average RT, and at the same time the greatest interference effect in case of neutral words, was observed in case of words from colours category ($M = 736$ ms, $SD = 110.513$ ms). The shortest average RT was noted in category of words related to sexuality and intimacy (without participation of children): $M = 656.49$ ms, $SD = 143.953$ ms. In this case an average RT was slightly below a value registered for neutral words. Relatively low average RT was also observed for words from a general category of words with sexual meaning ($M = 683.20$ ms, $SD = 132.875$ ms). Category of words related to child sexual abuse was on third place with respect to length of average RTs ($M = 701$ ms, $SD = 126.872$ ms).

To better illustrate differences between average RTs for words from sexual abuse category and words from sexuality and intimacy category, one used z-score standardisation. At the same time, to keep relation to an average RT noted for neutral stimuli category, a basis for standardisation was a mean value of RT and standard deviation for 3 indicated conditions (neutral and two sexual). Results of this operation are presented on Figure 2.

Table 3
Participants' average reaction time in milliseconds, standard deviations, and bias scores for emotional Stroop task

Word category	n	Mean [ms]	SD	Bias score [ms] (emo. – neutr.)
Neutral	21	666.13	84.4534	–
Colour	23	736.56	110.513	70.429
Negative emotions	22	711.51	176.242	45.3847
Child molesting	15	701.00	126.872	34.8769
Positive emotions	24	698.42	124.814	32.2904
Aggressive	23	691.61	129.650	25.4829
Sex/intimacy + child molesting	25	683.20	132.875	17.07
Sex/intimacy	10	656.49	143.953	–9.6331
Total	138	698.056	129.046	–

Positive bias scores indicate that the participant demonstrated a disadvantage in responding to words in that class compared to words in the neutral class.

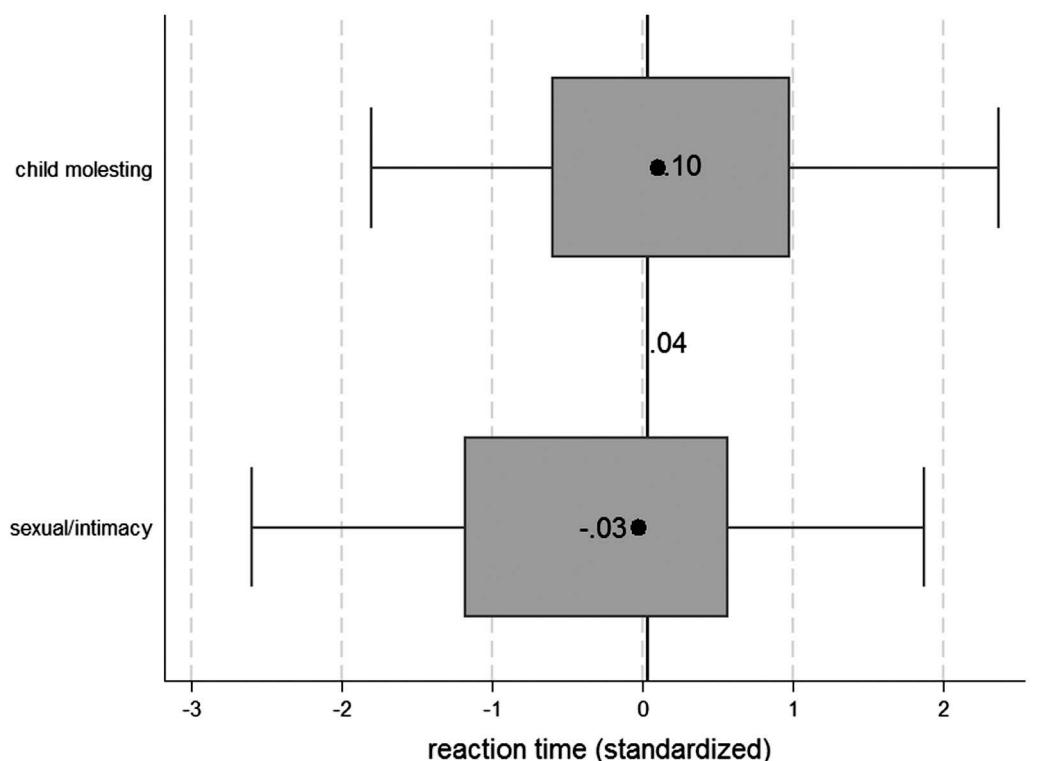


Fig. 2. Participant's reaction time (standardised) for child molesting and sex/intimacy word categories in comparison to neutral category. Dash line – median reaction time (standardised) for neutral category; black dots – median reaction times (standardised) for child molesting and sexuality/intimacy word categories.

As it was shown on Figure 2, over half of RTs noted for stimuli from sexuality/intimacy category were below median value for stimuli from neutral categories. Similarly, over half of stimuli from child sexual abuse category were over median for neutral stimuli. It is worthwhile to emphasise that the observed differences are rather slight. Difference between medians of both examined categories was only 0.13 of standard deviation ($SD = 112.518$ ms).

Because of the fact that the particular emotional conditions in Stroop task that were analysed were different with respect to the average number of letters in a single word, one conducted analysis of correlation of number of letters in a word with RT needed for its categorisation. Table 4 and Figure 3 presents the results.

Analysis of described relation proved a lack of relationship between a number of letters composing a word, and noted RT: $r(136) = 0.088, p = 0.340$.

8. Discussion of results

Analysis of described indicator in a way proposed by Smith and Waterman (2004), Price and Hanson (2007) and Price, Beech, Mitchell and Humpreys (2013) would not let increase diagnostic accuracy of undertaken decision.

Table 4
Mean number of letters for words in emotional Stroop task categories

Word category	<i>n</i>	Mean	<i>SD</i>
Neutral	21	5.80	1.549
Colour	23	7.52	2.563
Negative emotions	22	8.24	2.790
Positive emotions	24	7.12	2.046
Aggression	23	6.16	2.327
Child molesting	15	7.93	1.869
Sexuality/intimacy	10	6.55	2.147
Total	138	7.03	2.411

The examined person, manifesting sexual preferences in the form of paedophilia, presented the lowest RT interference in case of emotional condition, in which were words with sexual meaning. Argumentation, being a consequence of such result, one might formulate as follows: because of the fact that strength of interference in RT to stimuli related to sexuality and child sexual abuse was the lowest among all analysed emotional results, then there is no basis for a supposition that this type of stimuli is a source of a special distraction of attention for the examined person, and

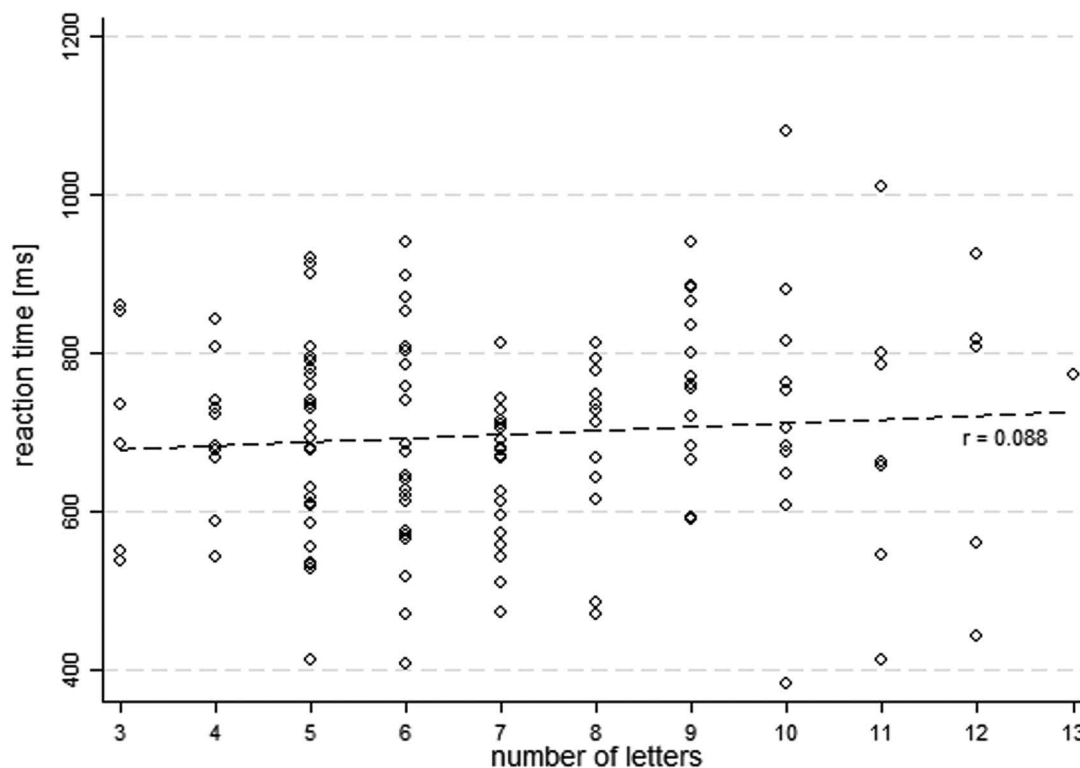


Fig. 3. Participant’s reaction time and number of letters for a word.

hence, that because of some reasons he is especially susceptible to them.

This type of reasoning provides an indirect argument for a lack of sexual tendencies for paedophilia in the examined person. In the light of facts that are known about the examined person, this conclusion will be incorrect.

Analysis of the indicator in a way proposed by teams of Ó Ciardha and Gormley (2012) and van Leeuwen et al. (2013) allows for formulating more accurate conclusions. Restricting the spectrum of interests to differences in average *RTs* between stimuli related to sexual abuse of minors and stimuli related to sexual behaviours that are within sexual norm, allows for formulation of the following conclusions: because the examined person needed longer reaction times to stimuli related to child sexual abuse than to stimuli related to sexual behaviours being within a sexual norm, it can be suspected that those stimuli are a source of special distraction for the examined person, significantly stronger than in a case of stimuli related to conventional sexual activities.

The reasoning presented above one may find as contributing to increase a diagnostic accuracy of issuing opinions on the presence of sexual preferences in the form of paedophilia in the examined person.

A slight difference between the two assessed conditions can be explained by sexual interest in adult women which was confirmed by the examined person (he consumes also this type of pornography and admits that he finds subtle and shy women attractive).

9. Problem of rejecting extreme values of *RT*

Conducted analysis made it possible to indicate problem of rejecting outliers of *RT*. In a situation when diagnostician has at his or her disposal a relatively low number of *RTs*, values that are significantly different very strongly influence the level of mean value for a given condition. On one hand, presence of this type of *RT* may be interpreted as an error resulting from a brief inattention or an ordinary mistake, etc. On the other hand, when a significantly different *RT* occurs, it may prove the special strength of the stimuli in attraction of examined person's attention. The discussed problem gets even more important when decision on leaving a given *RT* may settle the determination of distance between analysed mean values for particular categories of stimuli. It creates a possibility that a diagnostician may manipulate results, and, by rejecting or leaving consecutive *RTs*, s/he will be driven rather

by a wish to confirm his/her thesis than a need to obtain a reliable picture of situation.

Despite that the problems described above did not concern the presented situation (*RTs* in both categories related to sexuality created a relatively coherent picture), it is appropriate to formulate postulates diminishing the possibility of their occurrence.

10. Remarks on indicators of paedophilia based on reaction time

As indicate both studies discussed in this paper, and the author's experience coming from the above described experiment, one of the key characteristics of a potential tool supporting individual diagnosis on the basis of *RT* is taking into consideration the heterogeneity of population of persons manifesting sexual preferences in the form of paedophilia. One of conditions that should be met by the tool created for that purpose is a possibility of differentiation sexual preferences to at least two basic characteristics, namely age and sex. As it was shown by experiments described here, persons who sexually prefer young boys will not necessarily manifest this same pattern of reaction to stimuli presenting young girls. Due to differentiation of those two characteristics it also becomes possible to match results to examined persons who prefer both minors as well as sexually mature persons. On the basis of hitherto obtained results it seems to be legitimate to state that taking into account additional categories of emotional variants of Stroop task, such as aggression or positive emotions, seems to be a rather missed idea from the point of view of diagnosis of sexual preferences in the form of paedophilia.

Another remark concerns a number of stimuli and their repetitions presented in the consecutive conditions of Stroop task. To diminish uncertainty and arbitrariness of situations related to rejecting extreme *RTs*, it should be valuable to increase a number of presented stimuli. An important postulate seems to be also a multiple presentation of the same stimuli during the executed task. Thanks to that it is possible to assess if especially longer *RT* to a given stimulus is present each time, when this stimulus occurs, or if it is an example of a single incident.

11. Summary

Despite numerous promising research results in forensic sexology on use of indicators based on cognitive tasks, development of a tool supporting diagnostic

process, that would be based on them, still seems to be a very remote perspective. Attempts to use emotional Stroop task for that purpose are a good illustration of problems which have majority of researchers testing particular tasks and effects.

Among the most important objections towards the previous attempts are those concerning their accuracy. One should mention here insufficient and most often inconsequent use of external diagnostic criteria by their authors. Most often it concerns sentencing for sexual crimes against minors or self-definition of own sexual preferences as paedophilia. Still there is a lack of important information, among others on age of victims, preferences towards their sex, information if an examined person manifests only sexual preferences in the form of paedophilia, if s/he was ever diagnosed by a specialist sexologist, and so on. A clearly neglected aspect of those papers is also ecological validity, which in the discussed case requires arousal of strong motivation in examined persons to obtain the result they desire.

Despite the fact that author's attempt to use Stroop task for diagnosis of sexual preferences in the form of paedophilia described here seemed to be an accurate indicator, one cannot forget how small was the described difference in *RT*. From one side one can interpret this as an evidence for examined person's sexual preferences embracing conventional sexual behaviours, but at the same time it may be a manifestation of imperfection of the used procedure. The conducted research allowed for observing two practical problems related to use of this type of indicators. First of them concerns necessity to take into account stimuli that make it possible to differentiate not only age of preferred sexual objects, but also their sex while constructing them. Due to taking into account the mentioned reservation, it becomes possible to verify more detailed hypotheses concerning the examined person. Another problem concerned too small a number of repetitions used in particular categories of stimuli and necessity to present each of stimuli of a given type repeatedly. Application of the latter of remarks will clearly improve the decision process on a possible rejection of extreme cases, at the same time diminishing the risk of modelling results towards confirmation of own hypotheses by a diagnostician.

References

1. Banich, M. T., Crowley, T. J., Thompson, L. L., Jacobson, B. L., Liu, X., Raymond, K. M., Claus, E. D. (2007). Brain activation during the Stroop task in adolescents with severe substance and conduct problems: A pilot study. *Drug and Alcohol Dependence*, *90*, 175–182. DOI: (10.1016/j.drugalcdep.2007.03.009).
2. Beckett, R. C., Fisher, D., Mann, R. E., Thornton, D. (1997). Relapse prevention questionnaire and interview. (In) H. Eldridge (ed.), *Therapists guide for maintaining change: Relapse prevention manual for adult male perpetrators of child sexual abuse* (pp. 137–150). London: Sage.
3. Foa, E. B., Feske, U., Murdock, T. B., Kozak, M. J., McCarthy, P. R. (1991). Processing of threat related information in rape victims. *Journal of Abnormal Psychology*, *100*, 156–162. DOI: 10.1037/0021-843X.100.2.156.
4. Gotlib, I. H., McCann, C. D. (1984). Construct accessibility and depression: an examination of cognitive and affective factors. *Journal of Personality and Social Psychology*, *47*, 427–439. DOI: 10.1037/0022-3514.47.2.427.
5. Hope, D. S., Rapee, R. M., Heimberg, R. G., Dombek, M. J. (1990). Representations of the self in social phobia: vulnerability to social threat. *Cognitive Research Therapy*, *14*, 477–485.
6. MacLeod, C., Matthews, A., Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*, 15–20. DOI: 10.1007/BF01176208.
7. Ó Ciardha, C., Gormley, M. (2012). Using a pictorial-modified Stroop Task to explore the sexual interests of sexual offenders against children. *Sexual Abuse: A Journal of Research and Treatment*, *24*, 175–197. DOI: 10.1177/1079063211407079.
8. Pflugradt, D. M., Allen, B. P. (2010). An exploratory analysis of executive functioning for female sexual offenders: A comparison of characteristics across offense typologies. *Journal of Child Sexual Abuse*, *19*, 434–449. DOI: 10.1080/10538712.2010.495701.
9. Price, S. A., Beech, A., Mitchell, I. J., Humphreys, G. W. (2013). Measuring deviant sexual interest in adolescents using the emotional Stroop task. *Sexual Abuse: A Journal of Research and Treatment*, *26*, 450–471. DOI: 10.1177/1079063213495897.
10. Price, S. A., Beech, A. R., Mitchell, I. J., Humphreys, G. W. (2011). The promises and perils of the emotional Stroop task: A general review and considerations for use with forensic samples. *Journal of Sexual Aggression*, *18*(3), 253–268. DOI: 10.1080/13552600.2010.545149.
11. Price, S. A., Hanson, R. K. (2007). A modified Stroop task with sexual offenders: Replication of a study. *Journal of Sexual Aggression*, *13*, 203–216. DOI: 10.1080/13552600701785505.

12. Schmidt, A. F., Banse, R., Imhoff, R. (2015). Indirect measures in forensic contexts. (In) T. M. Ortner, F. J. R. van de Vijver (eds.), *Behavior-based assessment: Going beyond self-report in the personality, affective, motivation, and social domains* (pp. 173–194). Göttingen: Hogrefe.
13. Smith, P. (2009). Assessing sexual interest with the emotional Stroop test. (In) D. Thornton, D. R. Laws (eds.), *Cognitive approaches to the assessment of sexual interest in sexual offenders* (pp. 159–176). Chichester: Wiley.
14. Smith, P., Waterman, M. (2004). Processing bias for sexual material: The Emotional Stroop and sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 16, 163–171. DOI: 10.1023/B:SEBU.0000023064.54397.d6.
15. Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. DOI: 10.1037/h0054651.
16. Van Leeuwen, M., van Baaren, R., Chakhssi, F., Looenen, M. M., Lippman, M., Dijksterhuis, A. (2013). Assessment of implicit sexual associations in non-incarcerated pedophiles. *Archives of Sexual Behavior*, 42, 1501–1507. DOI: 10.1007/s10508-013-0094-0.

Appendix

Emotional Stroop words by category

Neutral	Aggression	Positive	Negative	Colour	Child molesting	Intimacy
Bag	Anger	Admire	Abandoned	Black	Molest	Touch
Ball	Annoyed	Affection	Abused	Bronze	Child	Naked
Chair	Argue	Affectionate	Afraid	Brown	Young	Buttocks
Cigarette	Assault	Amused	Aggression	Crimson	Boy	Girl
Diary	Crush	Beloved	Agony	Cyan	Caress	Flirt
Dog	Cut	Calm	Angry	Emerald	Helpless	Kiss
Door	Fight	Cheerful	Arrogant	Flesh	Schoolyard	Fondle
Eat	Guilt	Comfortable	Bad	Flesh	Innocent	Suck
Floor	Hate	Daring	Bitchy	Gold	Naive	Tempt
Football	Injure	Devotion	Crazy	Grey	Lure	Sex
Group	Kick	Euphoric	Cruelty	Grey	Pee pee	
Milk	Kill	Fond	Detest	Jade	Play	
Newspaper	Knife	Fond	Envy	Maroon	Dependent	
Oven	Punch	Funny	Fear	Mauve	Incest	
Painting	Punish	Glad	Frighten	Orange	Secret	
Pencil	Rage	Grateful	Fury	Orange		
Radio	Scream	Hope	Glum	Pink		
School	Shout	Joy	Greed	Pink		
Shoe	Slash	Lively	Hateful	Purple		
Shopping	Smash	Love	Lose	Purple		
Sofa	Tear	Peace	Misery	Scarlet		
Telephone	Temper	Protective	Pain	Silver		
Cigarette	Threaten	Proud	Spiteful	Tin		
Umbrella	Wound	Tolerant	Suspicious	Topaz		
Windy	Wound	Warm	Rage	White		

Data adapted from Price & Hanson (2007).

Corresponding author

Krzysztof Kasperek
 ul. Piwna 22/8
 PL 30-527 Kraków
 e-mail: krzysztof.kasperek@uj.edu.pl

WYKORZYSTANIE ZADANIA STROOPA W WERSJI EMOCJONALNEJ JAKO NARZĘDZIA POMOCNICZEGO W DIAGNOZIE PEDOFILNYCH PREFERENCJI SEKSUALNYCH¹

1. Wprowadzenie

Zadanie Stroopa (Stroop, 1935) zostało zaprojektowane jako jedno z zadań służących do pomiaru procesów przetwarzania informacji. Zadaniem osoby badanej jest poprawne kategoryzowanie kolorów, jakimi napisane są prezentowane jej słowa o różnym znaczeniu semantycznym. Czas reakcji pomiędzy wyświetleniem się danego słowa a przyporządkowaniem mu koloru przez badanego uznawany jest za wskaźnik interferencji zachodzącej pomiędzy znaczeniem semantycznym słowa a zautomatyzowanym procesem nazywania koloru. Zakłada się, że różnice w czasach reakcji potrzebnych na kategoryzację koloru prezentowanych kategorii słów są powodowane zróżnicowaniem ich ważności/znaczenia dla osoby badanej. Reakcje mogą polegać zarówno na naciskaniu przycisku odpowiadającego za dany kolor, jak i na głośnym wypowiedzeniu nazwy danego koloru.

Wśród najpopularniejszych modyfikacji zadania Stroopa P. Smith (2009) wymienia:

- słowa i kolory – badanemu prezentowane są różne kategorie słów, do których należy przyporządkować kolory, niezależnie od ich znaczenia semantycznego;
- słowa i obrazy – badanemu prezentowane są obrazy, w których osadzone są słowa. Zadaniem badanego jest kategoryzowanie obrazów, niezależnie od znaczenia semantycznego słów;
- dźwięki i słowa – zadaniem badanego jest klasyfikowanie prezentowanych dźwięków jako wysokie i niskie, niezależnie od znaczenia semantycznego dźwięku (np. słowo „niski” prezentowane wysokim tonem dźwięku);
- globalny-lokalny – zadaniem badanego jest kategoryzowanie globalnego (ogólnego) kształtu bodźców, niezależnie od sposobu ich prezentacji (np. cyfra pięć złożona z małych cyfr siedem).

Innym, bardzo często spotykanym typem modyfikacji, jest tzw. wersja emocjonalna (ang. emotional Stroop). Polega ona na dołączeniu do zadania dodatkowej zmiennej w postaci walencji emocjonalnej prezentowanych bodźców. Bodźce wymagające poprawnej kategoryzacji koloru miały semantyczne znaczenie związane z emocją

smutku, radości, gniewu etc. Opisywany rodzaj modyfikacji może zawierać także bodźce związane z dysfunkcjami przejawianymi przez osoby badane z grupy eksperymentalnej. Przykładowo w badaniu, w którym uczestniczą osoby uzależnione od alkoholu, jedna z grup bodźców bezpośrednio nawiązuje do alkoholu (np. słowo „kieliszek”, zdjęcie butelki piwa itd.).

Tak zmodyfikowana wersja zadania Stroopa była używana w badaniach z udziałem populacji osób cierpiących m.in. na zaburzenia lękowe (MacLeod, Matthews, Tata, 1986, za: Smith, 2009), depresję (Gotlib, McCann, 1984, za: Smith, 2009), PTSD (Foa, Feske, Murdock, Kozak, McCarthy, 1991, za: Smith, 2009), fobię społeczną (Hope, Rapee, Heimberg, Dombek, 1990, za: Smith, 2009) oraz uzależnienia od substancji (Banich i in., 2007, za: Price, Beech, Mitchel, Humphreys, 2013). W świetle wyników przedstawionych badań osoby cierpiące na wyżej wymienione dysfunkcje potrzebowały znacząco więcej czasu na kategoryzację koloru słów, które związane były z przejawianymi przez nich zaburzeniami, w porównaniu do osób z grup kontrolnych. Innymi słowy, bodźce pochodzące z kategorii problematycznych dla badanych populacji wymagały dłuższych *RT* (ang. Reaction Time) niż bodźce z kategorii neutralnych, w porównaniu do osób z grup kontrolnych (Price, Hanson, 2007).

2. Zadanie Stroopa w badaniach sprawców seksualnego wykorzystywania dzieci

Zaobserwowane w zadaniach poznawczych rezultaty wydłużenia reakcji na bodźce związane z przejawianymi dysfunkcjami stały się inspiracją dla badaczy sprawców przestępstw seksualnych. Upatrywano w tym zarówno szansę na zgłębienie mechanizmu popychającego do popełniania wspomnianych przestępstw, jak i możliwość wyłonienia obiektywnego, mniej inwazyjnego, niedającego się łatwo zmanipulować, tańszego w eksploatacji wskaźnika zaburzonych preferencji seksualnych. Za pracę, w której zapoczątkowano to podejście, można uznać artykuł P. Smitha i M. Watermana (2004), który ukazał się na łamach czasopisma *Sexual Abuse: A Journal of Research and Treatment*. Autorzy poddali procedurze zmodyfikowanego zadania Stroopa w wersji emocjonalnej sprawców przestępstw seksualnych przeciwko dzieciom. Jako grupę kontrolną dobrano sprawców przestępstw seksualnych przeciwko dorosłym kobietom, sprawców przestępstw agresywnych niemających charakteru seksualnego, sprawców przestępstw nieagresywnych oraz

¹ Niniejszy artykuł stanowi fragment niepublikowanej pracy magisterskiej pt. *Możliwości wykorzystania wskaźników opartych na czasie reakcji w diagnozie zaburzeń pedofilnych na przykładzie zadania Stroopa w wersji emocjonalnej* napisanej pod kierunkiem prof. dr hab. J. K. Gierowskiego, a obronionej w Instytucie Psychologii UJ.

studentów uniwersytetu w Leeds. Price i in. (2013), opisując najważniejsze z uzyskanych wtedy rezultatów, wskazują na:

- relatywnie dłuższy przeciętny *RT* w przypadku słów nacechowanych seksualnie wśród sprawców przestępstw seksualnych w porównaniu do pozostałych grup;
- relatywnie dłuższy przeciętny *RT* w przypadku słów z kategorii agresja u sprawców przestępstw agresywnych. Sprawcy zgwałceń potrzebowali w tym warunku relatywnie więcej czasu niż sprawcy przestępstw mniej agresywnych;
- brak istotnych różnic w prezentowanych *RT* pomiędzy sprawcami przestępstw seksualnych wobec dzieci i dorosłych.

Podsumowując swoje badania, Smith i Waterman (2004) stwierdzili, że zastosowane w zadaniu Stroopa słowa zbliżone treściowo do charakteru przestępstw badanych sprawców aktywują u nich, nieobecną u pozostałych badanych, sieć skojarzeń (ang. *cognitive associations*). Według autorów odzwierciedleniem tego właśnie procesu jest charakterystyczny wzorec relatywnie dłuższych *RT*.

Śmiała hipoteza oraz obiecujące rezultaty spowodowały prawdziwy wysyp badań nad wykorzystaniem różnego rodzaju zadań poznawczych w badaniu sprawców przestępstw seksualnych. Swoich kontynuacji doczekało się także zadanie Stroopa w wersji emocjonalnej. Jego replikacje w wykonaniu poszczególnych autorów z udziałem sprawców przestępstw seksualnych wobec dzieci zostały zestawione w tabeli 1².

Będąc częściowo replikacją opisanego eksperymentu Smitha i Watermana badanie Price i Hansona (2007) przyniosło podobne rezultaty. Przeciętne *RT*, potrzebne sprawcom zgwałceń na kategoryzację kolorów słów nacechowanych seksualnie, były wyraźnie dłuższe niż w warunku neutralnym w porównaniu do osób, które nie były skazane za żadne przestępstwa. Na szczególną uwagę zasługuje fakt, że autorzy wykorzystali w swoim eksperymencie zarówno zadanie Stroopa z zestawem słów zaproponowaną przez Smitha i Watermana oraz własny zestaw emocjonalnie nacechowanych kategorii. W przypadku zestawu drugiego nie odnotowano istotnych statystycznie różnic w zakresie czasu odpowiedzi. Autorzy tłumaczą to przyjętą „intuicyjną” metodą doboru zestawu słów. Zestawienie kategorii słów użytych w omawianych

² W sposób rozmyślny zrezygnowano z włączenia do listy badania Pflugradt i Allen (2010). Dotyczyło ono kobiet sprawczyń przestępstw seksualnych, z których wiele dopuściło się seksualnego wykorzystania dziecka, jednak użyte w badaniu zadanie Stroopa było zastosowane jedynie w wersji klasycznej, a jego celem była diagnoza sprawności kontroli poznawczej. W przypadku drugiego z badań, próbę przestępców seksualnych stanowili wyłącznie sprawcy młodociani, a podane na ich temat informacje nie pozwalały ustalić, czy znajdowali się wśród nich sprawcy wykorzystywania seksualnego dzieci.

badaniach oraz w pozostałych analizowanych eksperymentach zamieszczono w tabeli 2.

Innym ważnym badaniem z użyciem emocjonalnej wersji efektu Stroopa był eksperyment Ó Ciardha i Gormley (2012). Zadanie badanych polegało w nim na poprawnej kategoryzacji kolorów obrazów przedstawiających postacie obojga płci osób dorosłych, dzieci oraz kotów w warunku kontrolnym. Jest to jedyne spośród omawianych badań, w którym autorzy uwzględnili osoby o orientacji homoseksualnej zarówno wśród sprawców przestępstw seksualnych wobec dzieci, jak i wśród osób z grupy kontrolnej. Dzięki zastosowanemu zabiegowi możliwe stało się wydzielenie obrazów zgodnych z zainteresowaniami seksualnymi badanych (np. dorosłe kobiety). Jednym z najbardziej interesujących rezultatów otrzymanych w omawianym eksperymencie jest istotna statystycznie relatywna różnica w *RT* w przypadku obrazów dzieci o płci zgodnej z zainteresowaniami seksualnymi sprawcy przestępstwa w porównaniu do obrazów dzieci płci niezgodnej z zainteresowaniami sprawcy. Warto podkreślić fakt, że efekt ten zaobserwowano jedynie w przypadku sprawców otwarcie przyznających się do preferowania dzieci w kontaktach seksualnych.

W badaniu van Leeuwen i in. (2013) zadanie Stroopa zostało użyte w wersji słowa-obrazy. Osoby badane klasyfikowały wyświetlające się im wmontowane w obrazy słowa jako seksualne lub neutralne. Obrazy, podobnie jak w poprzednio opisanym badaniu, przedstawiały postacie obojga płci, zarówno dzieci, jak i dorosłych. W przedstawionym wariancie zadania Stroopa, nazywanym przez autora Testem Skojarzeń Obrazowych (*Picture Association Task*)³ uzyskano efekt odwrotny niż w pozostałych badaniach. Osoby samoidentyfikujące się z pedofilnymi zainteresowaniami seksualnymi uzyskiwały wyraźnie krótsze czasy w sytuacjach, gdy słowo nacechowane seksualnie było wkomponowane w obraz małego dziecka niż w pozostałych warunkach, w porównaniu do badanych z grupy kontrolnej. Warto zwrócić uwagę na fakt, że w przytoczonym badaniu analizy były prowadzone oddzielnie dla osób o hetero- i homoseksualnych pedofilnych preferencjach seksualnych.

Ostatnie z przytoczonych badań (Price, Beech, Mitchell, Humphreys, 2013), podobnie jak eksperyment Price i Hanson (2007), składało się z replikacji procedury Smitha i Watermana (2004) oraz nowego testowego zestawu słów. Jedną z jego najważniejszych cech był sposób, w jaki zostały dobrane słowa nacechowane seksualnie. Na podstawie odpowiedzi sprawców przestępstw seksualnych na jedno z pytań otwartych zawartych w *Relapse Prevention Questionnaire* (Beckett, Fisher, Mann, Thornton; 1997, za: Price, Beech, Mitchell, Humphreys,

³ Autor pracy, za przeglądem dokonany przez Schmidta, Banse i Imhoffa (2015), zdecydował się na klasyfikowanie użytego zadania jako zmodyfikowanej wersji zadania Stroopa.

2013), wyselekcjonowano określenia czynności seksualnych pojawiające się najczęściej w ich wypowiedziach. Analogicznie do badań Price i Hansona (2007), zadanie Stroopa, oparte na zestawie słów Smitha i Watermana, wykazało przeciętnie dłuższy *RT* potrzebny na udzielenie odpowiedzi w warunku seksualnym niż w warunku neutralnym w przypadku sprawców przestępstw seksualnych, w porównaniu do grupy osób niekaranych. Ten sam efekt zaobserwowano wśród sprawców przestępstw seksualnych, których ofiarami padały wyłącznie osoby dorosłe w porównaniu ze sprawcami przestępstw agresywnych. Wspomnianych różnic nie udało się natomiast wykazać pomiędzy ogółem sprawców przestępstw seksualnych a sprawcami przestępstw agresywnych.

3. Osoba badana

Osobą badaną był G. – 26-letni mężczyzna, absolwent technikum informatycznego. G. po ukończeniu szkoły średniej podejmował nieudane próby studiów oraz pracy fizycznej. Z każdej z prób rezygnował ze względu na trudności w zaadaptowaniu się w grupie społecznej i odczuwany w związku tym silny lęk. W swoim życiu G. nie budował bliskich więzi z rówieśnikami. Nie był również nigdy w żadnym związku intymnym. G. uczestniczył w terapii indywidualnej, do której zgłosił się dobrowolnie, ze względu na przejawiane zainteresowanie seksualne dziećmi. Pedofilne preferencje seksualne zaobserwował u siebie w okresie adolescencji w wieku ok. 17–18 lat. Mniej więcej od tego czasu pojawiły się u niego fantazje na temat utrzymywania kontaktów seksualnych z chłopcami w wieku 10–12 lat. Poza fantazjowaniem, G. użytkował również materiały pornograficzne prezentujące chłopców we wspomnianym wieku. W rozmowie z terapeutą G. zaprzeczył, aby kiedykolwiek podejmował kontakty seksualne z osobami znajdującymi się w wieku ochronnym. Nie podejmował również żadnych kontaktów seksualnych z osobami dojrzałymi płciowo. Poza pornografią pedofilską, G. użytkował także materiały pornograficzne z udziałem osób dorosłych. Jako atrakcyjne seksualnie opisywał delikatne, nieśmiałe kobiety.

Po rozmowie z terapeutą i krótkim zapoznaniu go z celem badania, G. zgodził się na udział w nim. Podczas kontaktu z eksperymentatorem, po przypomnieniu celu badania i krótkim opisie procedury, G. podtrzymał swoją chęć uczestnictwa.

4. Procedura badawcza

Zastosowana procedura badawcza została oparta o eksperyment Smitha i Watermana (2004), w której kategorię słów nacechowanych seksualnie zastąpiono kategorią słów „molestowanie seksualne dzieci” z ekspe-

rymentu Price i Hanson (2007). Finalny zestaw zawierał 150 słów pogrupowanych w siedem kategorii tematycznych: neutralne, pozytywne, związane z agresją, emocje pozytywne, emocje negatywne, kolory oraz słowa związane z molestowaniem seksualnym dzieci. Słowa użyte w procedurze wraz z ich oryginalnymi wersją znajdują się w aneksie.

Procedura badawcza została zaprogramowana w programie DMDX na laptopie marki Epson E-Machines model E525z, ekran 15,60”, procesor 2 GHz, system operacyjny Windows XP. Reakcje badanego polegały na naciśnięciu jednego z klawiszy (A, S, K, L), odpowiadającemu kolorowi wyświetlanego słowa (czerwony, zielony, żółty lub niebieski). Aby ułatwić badanemu udzielanie poprawnych reakcji, na klawisze naklejone zostały kolorowe znaczniki odpowiadające barwie przyporządkowanej do klawisza. Warto zwrócić uwagę, że procedura różni się od dwóch oryginalnych eksperymentów, gdzie zadanie osób badanych polegało na głośnym wymówieniu nazwy koloru prezentowanego słowa.

Procedura rozpoczęła się od instrukcji, po zapoznaniu z którą badany odbył sesję treningową liczącą 29 prób. W trakcie treningu otrzymywał on informacje zwrotne dotyczące poprawności lub braku poprawności swoich reakcji. Każda z prób rozpoczynała się od wyświetlenia przez 500 ms punktu fiksacji „X” w centralnym punkcie ekranu, po którym wyświetlało się losowo jedno ze 150 słów. Każde z nich mogło wyświetlić się tylko raz w trakcie procedury. Komunikat informujący o zakończeniu zadania pojawiał się po reakcji badanego na ostatnie z wyświetlonych słów.

5. Analiza wyników

Podobnie jak w oryginalnych badaniach Smitha i Watermana (2004) oraz Price i Hanson (2007), w przeprowadzonym badaniu przedmiotem analiz były różnice pod względem *RT* w warunkach emocjonalnych w stosunku do *RT* w warunkach neutralnych.

Dodatkowa modyfikacja, nieobecna w opisywanych eksperymentach, polegała na wydzieleniu z kategorii słów związanych z molestowaniem seksualnym dzieci dziesięciu słów dotyczących ogólnie zachowań seksualnych i intymności, niezwiązanych bezpośrednio z dziećmi. Podział ten został przeprowadzony zgodnie z sugestią autorów opisywanego eksperymentu (Price, Hanson, 2007).

6. Wylączenie z analiz

Pomimo uprzedniego testowania procedury zadania Stroopa, w oryginalnej wersji zadania znalazły się drobne błędy polegające na nieprawidłowym wyświetleniu się pięciu słów, wśród których dwa pochodziły z kategorii

agresja, dwa z kategorii emocje negatywne oraz jedno z kategorii kolory. W przypadku czterech z nich błędy polegały na niewyświetleniu się znaków specjalnych (ą, ć, ł). Ostatnie z problematycznych słów wyświetliło się rozdzielone spacją („agre_sywny”). W celu zniwelowania wpływu błędnego sposobu wyświetlania słów na *RT* badanego zostały one wyłączone z dalszych analiz.

Podobnie jak w oryginalnych badaniach, czasy błędnych reakcji (niepoprawne skategoryzowanie koloru) nie były włączane do analiz. Z analizy zostały usunięte *RT* wyraźnie odstające od pozostałych (ang. outliers). Związana z tym problematyka oraz schemat postępowania zostanie dalej szerzej omówiony.

7. Wyniki badania

W trakcie realizacji procedury badany popełnił jeden błąd w kategoryzacji kolorów w słowie z kategorii neutralne. Wynik został usunięty z dalszych analiz. Wykres rozrzutu *RT* badanego w podziale na kategorie słów został zaprezentowany na rycinie 1.

Na podstawie analizy przedstawionego wykresu zdecydowano się na usunięcie *RT* znajdujących się poza 5 i 95 percentylem, tj.:

- 3 obserwacje z kategorii neutralne;
- 1 obserwacja z kategorii emocje pozytywne;
- 1 obserwacja z kategorii emocje negatywne;
- 1 obserwacja z kategorii kolory.

Usunięte obserwacje stanowiły 8% wszystkich zarejestrowanych *RT*. Statystyki opisowe tak przygotowanych *RT* przedstawiono w tabeli 3. Kategorię słów, zaczerpniętą z badania Price i Hanson (2007), przedstawiono zarówno w postaci zbiorczej, jak i w rozbięciu na opisywane wyżej podkategorie związane z czynnościami seksualnymi dotyczącymi dzieci oraz ogólnie rozumianą seksualnością i intymnością (bez udziału dzieci).

Najdłuższy przeciętny *RT*, a zarazem największy efekt interferencji w stosunku słów neutralnych, został odnotowany w przypadku słów z kategorii kolory ($M = 736$ ms, $SD = 110,513$ ms). Najkrótszy przeciętny *RT* odnotowano w kategorii słów związanych z seksualnością i intymnością (bez udziału dzieci): $M = 656,49$ ms, $SD = 143,953$ ms. Przeciętny *RT* znalazł się w tym przypadku nieznacznie poniżej wartości odnotowanej dla słów neutralnych. Relatywnie niski przeciętny *RT* odnotowano także dla słów z ogólnie potraktowanej kategorii słów nacechowanych seksualnie ($M = 683,20$ ms, $SD = 132,875$ ms). Kategoria słów związanych z molestowaniem seksualnym dzieci znalazła się na trzecim miejscu pod względem długości przeciętnych *RT* ($M = 701$ ms, $SD = 126,872$ ms).

Aby lepiej zobrazować różnice pomiędzy przeciętnymi *RT* dla słów z kategorii molestowanie seksualne a słowami z kategorii seksualność i intymność, postu-

żono się standaryzacją typu *z*. Aby zachować przy tym odniesienie do przeciętnego *RT* odnotowanego dla kategorii bodźców neutralnych, podstawą do standaryzacji była średnia wartość *RT* oraz odchylenie standardowe dla 3 odanych warunków (neutralnego oraz dwóch seksualnych). Rezultaty tego zabiegu zobrazowano na rycinie 2.

Jak wykazano na rycinie 2, ponad połowa *RT* odnotowanych dla bodźców z kategorii seksualność/intymność, znalazła się poniżej wartości mediany dla bodźców z kategorii neutralnych. Analogicznie, ponad połowa bodźców z kategorii molestowanie seksualne dzieci znalazła się powyżej mediany dla bodźców neutralnych. Warto przy tym zaznaczyć, że obserwowane różnice są raczej niewielkie. Różnica pomiędzy medianami obu badanych kategorii wynosi jedynie 0,13 odchylenia standardowego ($SD = 112,518$ ms).

Ze względu na fakt, że analizowane poszczególne warunki emocjonalne w zadaniu Stroopa różniły się od siebie pod względem przeciętnej liczby liter przypadających na jedno słowo, przeprowadzono analizę korelacji liczby znaków w słowie i *RT* wymaganego na jego kategoryzację. Rezultaty analiz przedstawiono w tabeli 4 oraz na rycinie 3.

Analiza opisywanej zależności wykazała brak związku pomiędzy liczbą liter, z których składa się dany wyraz, a odnotowywanym *RT*: $r(136) = 0,088$, $p = 0,340$.

8. Dyskusja wyników

Analiza opisywanego wskaźnika w sposób zaproponowany przez Smitha i Watermana (2004), Price i Hanson (2007) oraz Price, Beech, Mitchell i Humphreys (2013), nie pozwoliłaby na zwiększenie trafności diagnostycznej podejmowanej decyzji.

Badana osoba przejawiająca pedofilne preferencje seksualne wykazała najmniejszą interferencję *RT* w przypadku warunku emocjonalnego, w którym znalazły się słowa nacechowane seksualnie. Argumentację, będącą następstwem takiego wyniku, można by sformułować następująco: ze względu na fakt, że wielkość interferencji w *RT* na bodźce związane z seksualnością oraz molestowaniem seksualnym dzieci była najniższa spośród wszystkich analizowanych warunków emocjonalnych, nie ma podstaw, aby przypuszczać, że bodźce tego typu stanowią dla badanego źródło szczególnej dystrykcji uwagi, a co za tym idzie, że jest on z jakichś względów szczególnie na nie wrażliwy.

Tego typu rozumowanie dostarcza pośredniego argumentu na rzecz braku pedofilnych preferencji seksualnych u badanego. W świetle znanych faktów na temat badanego będzie to więc wniosek błędny.

Analiza wskaźnika w sposób zaproponowany przez zespoły Ó Ciardha i Gormley (2012) oraz van Leeuwen i in. (2013) pozwala na sformułowanie bardziej trafnych

wniosków. Zawężenie spektrum zainteresowań do różnic w przeciętnych *RT* pomiędzy bodźcami związanymi z seksualnym molestowaniem nieletnich a bodźcami związanymi z zachowaniami seksualnymi mieszczącymi się w normie seksuologicznej, pozwala sformułować następujące wnioski: ponieważ osoba badana potrzebowała dłuższych czasów reakcji na bodźce związane z seksualnym molestowaniem dzieci niż na bodźce związane z zachowaniami seksualnymi mieszczącymi się w normie seksuologicznej, zachodzi podejrzenie, że bodźce te stanowią dla badanego źródło szczególnej dystrakcji, wyraźnie większej, niż w przypadku bodźców związanych z konwencjonalnymi aktywnościami seksualnymi.

Przedstawione wyżej wnioskowanie można uznać za przyczyniające się do zwiększenia trafności diagnostycznej orzekania na temat obecności pedofilnych preferencji seksualnych u opiniowanego.

Niewielką różnicę pomiędzy dwoma ocenianymi warunkami można tłumaczyć potwierdzonym przez badanego zainteresowaniem seksualnym kobietami dojrzalymi płciowo (badany użytkuje także ten rodzaj pornografii oraz zgłasza, że pociągają go seksualnie delikatnie i nieśmiałe kobiety).

9. Problem odrzucania skrajnych wartości *RT*

Przeprowadzona analiza pozwoliła na zasygnalizowanie problemu odrzucania skrajnych wartości *RT*. W sytuacji, gdy do dyspozycji diagnosty pozostaje stosunkowo niewielka liczba *RT*, pojawiające się wartości odstające bardzo wyraźnie oddziałują na wielkość średniej dla danego warunku. Z jednej strony obecność tego rodzaju *RT* można interpretować jako błąd wynikający z chwilowego rozproszenia uwagi lub zwykłej pomyłki itd. Z drugiej strony pojawienie się wyraźnie odstającego *RT* może świadczyć o wyjątkowej sile, z jaką użyty bodziec ściągnął uwagę osoby badanej. Omawiany problem nabiera dodatkowej wagi w momencie, gdy decyzja o pozostawieniu danego *RT* może okazać się przesądzającą dla ustalenia odległości pomiędzy analizowanymi średnimi dla poszczególnych kategorii bodźców. Stawarza to możliwość manipulacji wynikami ze strony prowadzącego badanie diagnosty, który, odrzucając lub pozostawiając kolejne *RT*, będzie kierował się raczej chęcią potwierdzenia swojej tezy niż potrzebą uzyskania rzetelnego obrazu sytuacji.

Pomimo tego, że opisywane wyżej problemy nie dotyczyły przedstawionej sytuacji (*RT* w obu kategoriach związanych z seksualnością tworzyły relatywnie spójny obraz), właściwe wydaje się sformułowanie postulatów zmniejszających prawdopodobieństwo ich wystąpienia.

10. Uwagi dotyczące wskaźników pedofilnych preferencji seksualnych opartych na czasie reakcji

Jak pokazują zarówno analizowane w pracy badania, jak i doświadczenie autora płynące z przeprowadzenia opisanego wyżej eksperymentu, jedną z cech kluczowych potencjalnego narzędzia wspomagającego diagnozę indywidualną w oparciu o *RT* jest uwzględnienie heterogeniczności populacji osób przejawiających pedofilne preferencje seksualne. Jednym z wymogów, jakie powinno spełniać wytworzone w tym celu narzędzie, jest rozróżnienie preferencji seksualnych przynajmniej ze względu na dwie podstawowe charakterystyki, tj. wiek oraz płeć. Jak pokazały opisywane w pracy eksperymenty, osoby wykazujące preferujące seksualne wobec młodych chłopców niekoniecznie będą wykazywały ten sam wzór reakcji wobec bodźców przedstawiających młode dziewczęta. Dzięki rozróżnieniu tych dwóch cech możliwe staje się również dopasowanie wyników do opiniowanych preferujących zarówno osoby w wieku ochronnym, jak i dojrzałe seksualnie. Na podstawie dotychczas uzyskiwanych rezultatów uprawnione wydaje się stwierdzenie, że uwzględnianie dodatkowych kategorii emocjonalnych wariantów zadania Stroopa, takich jak agresja czy pozytywne emocje, wydaje się pomysłem raczej chybionym z punktu widzenia diagnozy pedofilnych preferencji seksualnych.

Druga uwaga dotyczy liczby bodźców oraz ich powtórzeń prezentowanych w ramach poszczególnych warunków zadania Stroopa. W celu zmniejszenia niepewności oraz arbitralności sytuacji związanych z usuwaniem *RT* skrajnie odstających od pozostałych, wartościowe winno być zwiększenie liczby prezentowanych bodźców. Ważnym postulatem wydaje się również kilkakrotne powtarzanie tych samych bodźców w ramach wykonywanego zadania. Dzięki temu możliwe staje się dokonanie oceny, czy szczególnie wydłużony *RT* na dany bodziec występuje za każdym razem, gdy ten się pojawia, czy też jest on przykładem jednorazowego przypadku.

11. Podsumowanie

Pomimo wielu obiecujących rezultatów badań prowadzonych nad wykorzystaniem wskaźników opierających się na zadaniach poznawczych w seksuologii sądowej, opracowanie bazującego na nich narzędzia wspomagającego proces diagnostyczny wydaje się ciągle bardzo odległą perspektywą. Próby wykorzystania w tym celu zadania Stroopa w wersji emocjonalnej stanowią dobrą ilustrację problemów, z którymi boryka się większość badaczy testujących poszczególne zadania i efekty.

Do najważniejszych zastrzeżeń pod adresem dotychczas podejmowanych prób należą zarzuty pod wzglę-

dem ich trafności. Należy tu wymienić niewystarczające i najczęściej niekonsekwentne stosowanie przez autorów zewnętrznego kryterium diagnostycznego. Najczęściej dotyczy ono skazania za przestępstwa seksualne wobec osób w wieku ochronnym lub samookreślenie swoich preferencji seksualnych jako pedofilnych. Brakuje przy tym ważnych informacji, m.in. na temat wieku ofiar, preferencji wobec ich płci, wiadomości o tym, czy badany przejawia wyłącznie pedofilne zainteresowania seksualne, czy był kiedyś diagnozowany przez specjalistę seksuologa itd. Wyraźnie zaniedbanym aspektem prac jest także trafność ekologiczna, która w omawianym przypadku wymaga wzbudzenia silnej motywacji w osobach badanych pod względem uzyskania pożądanego przez nie rezultatu.

Pomimo tego, że opisana w pracy autorska próba zastosowania zadania Stroopa do diagnozy pedofilnych zainteresowań seksualnych okazała się trafnym wskaźnikiem, nie należy zapominać, jak niewielka była opisywana różnica w *RT*. Z jednej strony można interpretować to jako dowód na występowanie u osoby badanej

preferencji seksualnych obejmujących konwencjonalne zachowania seksualne, choć równie dobrze może być to przejaw wadliwości zastosowanej procedury. Przeprowadzone badanie pozwoliło zauważyć dwa praktyczne problemy związane z korzystaniem z tego typu wskaźników. Pierwszy z nich dotyczy konieczności uwzględnienia w ich konstrukcji bodźców pozwalających na różnicowanie nie tylko wieku preferowanych obiektów seksualnych, ale również ich płci. Dzięki uwzględnieniu wspomnianego zastrzeżenia, możliwe staje się weryfikowanie bardziej szczegółowych hipotez dotyczących opiniowanego. Drugi z problemów dotyczył zbyt małej liczby powtórzeń stosowanych w poszczególnych kategoriach bodźców oraz konieczności kilkukrotnego wyświetlania każdego z bodźców danego typu. Zastosowanie się do drugiej z uwag wyraźnie usprawni proces podejmowania decyzji nt. ewentualnego odrzucania przypadków odstających, zmniejszając jednocześnie ryzyko modelowania wyników w kierunku potwierdzenia własnych hipotez przez diagnostę.

Aneks

Słowa wykorzystane w zastosowanym zadaniu Stroopa

Neutralne	Agresja	Emocje pozytywne	Emocje negatywne	Kolory	Molestowanie seksualne dzieci	Seks/intymność
Drzwi	Wściekłość	Oddanie	Opuszczony	Fioletowy	Molestować	Dotykać
Grupa	Gniew	Uczucie	Sponiewierać	Szary	Dziecko	Goły
Krzesło	Łza	Podziw	Lęk	Błady	Młody	Pośladki
Telefon	Napaść	Euforia	Agresywny	Różowy	Chłopczyk	Dziewczyna
Pies	Kopniak	Czuły	Umieranie	Pomarańczowy	Troskliwy	Flirt
Kot	Krzyk	Wdzięczny	Rozzłoszczony	Szkarłatny	Bezbronny	Całus
Sofa	Cios	Tolerancyjny	Arogancja	Kasztanowy	Podstawówka	Pieścić
Torba	Nienawiść	Uczuciowy	Zły	Karmazynowy	Niewinny	Ssać
Dziennik	Kłótnia	Rozbawiony	Wredny	Biały	Naiwny	Kusić
Gazeta	Porywczy	Miłość	Szaleństwo	Czarny	Zwabić	Seks
Jedzenie	Walka	Radość	Okrutnie	cyjan	Siusiak	
Piekarnik	Zabić	Duma	Obrzydliwy	Brązowy	Zabawa	
Podłoga	Karać	Miły	Zazdrosny	Cynowy	Zależny	
Zakupy	Wkurzony	Zabawny	Strach	Brąz	Kazirodztwo	
Parasol	Winny	Zadowolony	Zalękniony	Fioletkowy	Tajemnica	
Wietrzny	Krzyczeć	Wygodny	Zaciekły	Złoty		
Radio	Rozbić	Ulubiony	Ponury	Srebrny		
Obraz	Pociąć	Spokojny	Cheiw	Jadeitowy		
Mleko	Bić	Pokój	Nienawistny	Topazowy		
Biuro	Ciąć	Śmiały	Złośliwy	Szmaragdowy		
Piłka	Rana	Wesoły	Podejrzliwy	Fioletowy		
Ołówek	Okaleczyć	Ciepły	Rozgniewany	Szary		
Papieros	Grozić	Opiekuńczy	Ból	Błady		
Hokej	Nóż	Nadzieja	Strata	Różowy		
But	Spoliczkować	Żwawy	Niedola	Pomarańczowy		

Źródło: opracowanie własne na podstawie Price i Hanson (2007).